As for any negative effects of using the PTC, an additional eleven percent of students in the PTC Group believed that their pronunciation was *worse than other students* after practicing with the PTC. This negative statistic is worrisome since it is desired that all students feel a sense of improvement at the end of the term. Concerning why this rather negative assessment occurred, perhaps these students really are worse than other students. If one were to employ a Bell curve to student speaking ability, the probability that one group performs worse than the norm would certainly exist. This is an area that requires further research and analysis. If a case study were conducted where the individual's pre- and post-test scores and self-assessments were recorded, the researcher could gain greater understanding of this response.

In conclusion, after conducting this research for one term, it appears that the PTC had the greatest benefit when used in a one-on-one situation rather than sharing in small groups, and when students had an opportunity to practice with it for at least thirty minutes two times per week. Under these circumstances, improvement in both test scores and self-assessment of speaking and pronunciation ability were noted. Although the results of this research are not conclusive as to the benefits or drawbacks of the PTC as a learning aid, none-the-less, the study was worthwhile exploratory teaching wherein the teacher is involved in an "...ongoing process of working constantly for deeper understanding and increasing effectiveness in the classroom" (Nunan, 18).

Works Cited

Babbie, Earl. *The Practice of Social Research*. Belmont: Wadsworth Publishing, 1986.

Brookfield, Stephen D. The Skillful Teacher. San Francisco: Jossey-Bass, 1990.

Nunan, David. "Language Teaching and Result." *Classroom Teachers and Classroom Research*. Ed. Dale T. Griffin and David Nunan. Tokyo: The Japan Association for Language Teaching, 1997. 13-21.

Tanaka, Rebecca. "Automatic Speech Recognition and Language Learning." *The Journal of Wayo Women's University* 40 (March 2000): 53-62.

(人文学部国際社会学科助教授)

and some lower scores than predicted. For the 2X Week Group, actual scores matched predictions most consistently.

The results of the regression analysis indicate that use of the PTC alone did not produce a strong influence on test scores.

Conclusion

Although the preceding analysis is incomplete due to unavailability of pre-test ability scores, the original hypothesis, that study with the PTC helps students to improve their English speaking ability, remains uncertain.

The results would have been more conclusive had the research design included a pre-test or some other determination of student language ability at the beginning of the project. Additionally, if a twice-per-week small class that practiced the *Let's Go To New York* conversations using the L/I method had been included in the study, a more adequate comparison of test results could have been made. Finally, if the research design had designated code numbers for all participants' responses and test scores, more refined data would have been available to check the impact of other outside factors such as travel abroad, study at a language school, use of English outside of class, etc. Therefore, it is not possible to say conclusively what affect the PTC had on test scores.

On the other hand, when the participants' perceptions of their own language abilities and self-confidence are considered, the PTC had a positive impact. This is significant since, as Brookfield (1990) noted, to perform well, students need to talk frequently and confidently. Language becomes inextricably bound up with self-esteem. Therefore, a positive self-impression about the method of study and the results is an important factor in student learning.

At the end of the research period, all groups thought their speaking and pronunciation abilities had improved because of using the PTC. The PTC Group and the 2X Week Group indicated the greatest perception of improvement. Thirty-three percent of the students who practiced with the PTC thought their speaking was *much better than other students*, thirty-three percent thought it was *better than other students*, thirty-three percent thought they were the *same*, and none of them thought they were *worse or much worse than other students*.

used the PTC, not between the L/I Group and the PTC Group. This indicates that some other factor influenced scores.

PTC Group compared with L/I Group	0.0334133
2X Week Group compared with L/I Group	0.0007885
PTC Group compared with 2X Week Group	0.18627647

One final test, a regression analysis, was calculated to determine if use of the PTC created a causal effect on test scores. In setting up the data tables for the regression analysis, test scores for all students were entered as the constant variable Y. Another table of how many times students had practiced with the PTC was created for variable X. All three groups of students were included.

The most revealing aspect of the regression analysis was the fit line plot represented by Chart 3 on which predicted test scores and actual test scores were plotted for all three groups. If the original hypothesis were valid and study using the PTC did increase test scores, the plots for all predictions and actual scores would line up together. However, for the L/I Group, the fit line plot revealed that in only five cases did predictions match actual test scores, in two cases actual scores were one notch higher than actual scores and in three cases, actual scores were two notches higher than predictions. The remaining actual scores showed great variance from predictions. When analyzing the plotted scores for the PTC group, seven cases showed strong variance from predictions–some higher scores

Chart 3 : Regression Analysis Fit Line Plot

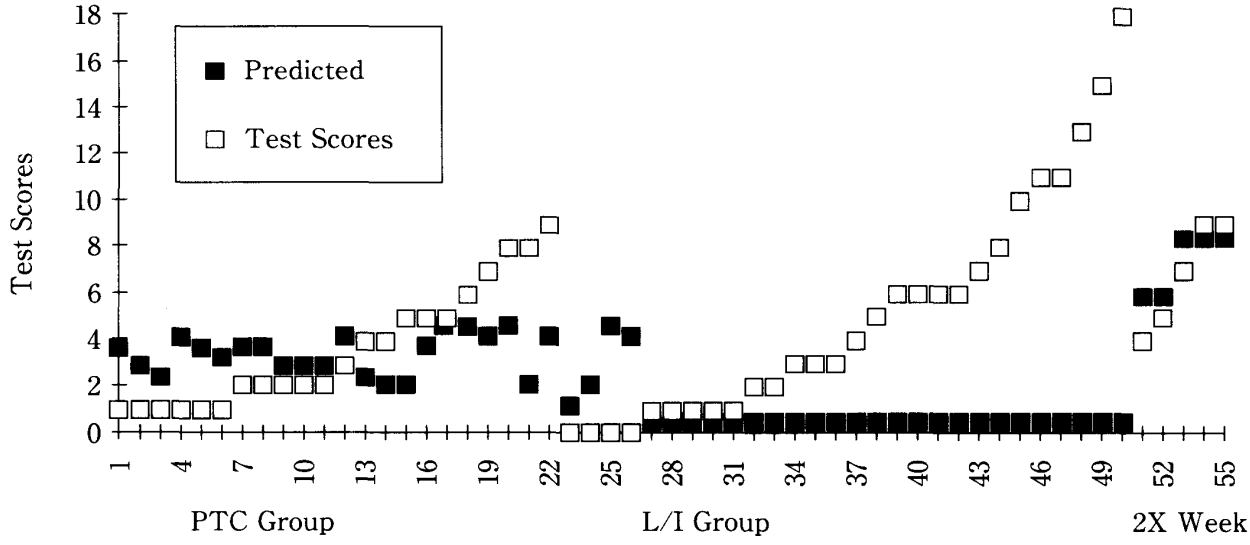


Table 6 : ANOVA Single-factor Results for All Tests, All Groups

	Count	Sum	Mean	Variance
PTC Group Machine	26	80	3.077	7.593
L/I Group Machine	26	144	5.538	24.184
2X Week Group Machine	5	32	6.400	3.800
PTC Group Teacher	28	63	2.250	1.676
L/I Group Teacher	28	146	5.214	10.397
2X Week Group Teacher	7	13	1.857	0.810
Source of Variation	SS	df	MS	F
Between Groups	271.638	5	54.438	5.450
Within Groups	1136.329	114	9.968	
Total	1407.967	119		

P-value 0.000155751, F crit 2.29390906

Table 7 : ANOVA Single-factor for PTC and L/I Group Machine Tests

	Count	Sum	Mean	Variance
PTC Group	26	80	3.077	7.593
L/I Group	26	144	5.538	24.184
Source of Variation	SS	df	MS	F
Between Groups	78.769	1	78.769	4.983
Within Groups	790.307	50	15.806	
Total	869.077	51		

P-value 0.30099933, F crit 4.03431955

L/I Group. The results of this test shown in Table 7 reveal an f-factor that is still significant at 4.983.

Subsequently, f-tests were conducted. Concerning the f-tests, if the PTC had contributed to a large portion of the variance, the variances between both the two groups who used the PTC and the L/I Group would be high. As we can see from the f-test values, the opposite results were obtained. The largest f-factor was between the two groups who

Table 5 : Correlation of Test Scores

Test Scores	PTC Machine	L/I Machine	2X Wk Machine	PTC Teacher	L/I Teacher	2X Wk Teacher
PTC Machine	1					
L/I Machine	0.136	1				
2X Machine	0.595	0.333	1			
PTC Teacher	0.242	0.908	0.494	1		
L/I Teacher	0.309	0.232	0.456	0.369	1	
2X Teacher	0.222	0.012	0.045	0.169	0.133	1

correct pronunciation and intonation. I noticed an additional sense of striving for accuracy that was not present in the L/I Group. This element may contribute to greater ease of understanding by a native speaker.

Statistical analyses for social science research contained in Babbie (1986) were conducted. Correlation of the data sets in Table 5 shows the strongest link is between L/I machine test scores and 2X Week groups with weak correlations among all other groups. Thus, a strong correlation between using the PTC equipment and test scores does not exist. These results suggest the 2X Week Group's stronger test performance was influenced by an outside factor, perhaps the amount of study rather than the method of study. Without a control group, it is impossible to draw solid conclusions.

In the ANOVA single-factor test that was conducted, if the PTC had positively affected test scores, there would be a low variance from the mean of test scores for the groups who used the PTC and high variance for those who did not. As we can see from the results of the ANOVA single-factor test contained in Table 6, variances for the PTC Group Machine Test and the 2X Week Group Machine Test are much lower than the L/I Group Machine Test, which points to a positive influence. An f-factor of 5.450 indicates a significant impact. Since the variation within groups is also high, it is possible that the high variance in the L/I Group is caused by differences in language ability among class members rather than by method of instruction.

In an attempt to get a clearer comparison of the effect of the PTC on test scores, a second ANOVA test was run using only the machine test scores for the PTC Group and the

choose to either replicate the PTC conversations or have original conversations. Table 4 contains the results of these tests.

In looking at the average, median, and mode machine test scores for the three groups, we see that the PTC Group scored lower than the other two groups while the 2X Week Group scored higher. There are several possible explanations for these results. First, because the PTC Group members shared the machines among four or five students, each individual had less practice time than the other two groups who were either practicing the materials in chorus with the teacher, with a partner, or for twenty to thirty minutes with the PTC individually. In the PTC Group, practice time with the machine was usually seven to eight minutes per person per class. Although students were supposed to practice the conversations individually during the thirty-minute study period while waiting their turn for the machine, in actuality, this seldom occurred, despite teacher encouragement.

A second possible explanation of the high scores by the 2X Week Group is that they not only practiced longer with the equipment, but also they had two class periods to interact with the materials each week. Since there was not a control group where students practiced the L/I method during an additional class each week, it is not possible to determine clearly whether it was the additional time with the equipment, the additional exposure to the materials or some other factor that created the higher scores. Student comments obtained from the end of term questionnaires, however, indicated that it was easier to study in the smaller class than in the larger classes. Even though other students in the larger class could not clearly hear what the speaker was saying, many students expressed that they experienced a sense of inhibition when speaking near many other students in the classroom.

The results of the teacher-monitored tests initially show that the 2X Week Group's average scores were lower than either of the other groups, although the mean and mode were identical. The explanation for this result is that the groups with the lower scores actually engaged in longer, more detailed conversations than their counterparts who received scores for *more* conversations. For the student's actual grades, this factor was taken into account and grades were adjusted accordingly.

Subjectively, as the monitoring teacher, I did not detect that the PTC created great improvement in the content of the conversations, but students were much more aware of

improvement. Table 3 shows the t-test results.

Table 3 : T-Tests Before and After Study

Confidence Level .05	Speaking Before	Speaking After	Pronunciation Before	Pronunciation After
df	39	39	45	45
t	1.05207721	3.44923053	.063052083	1.23618023

Results and Analysis of Test Scores

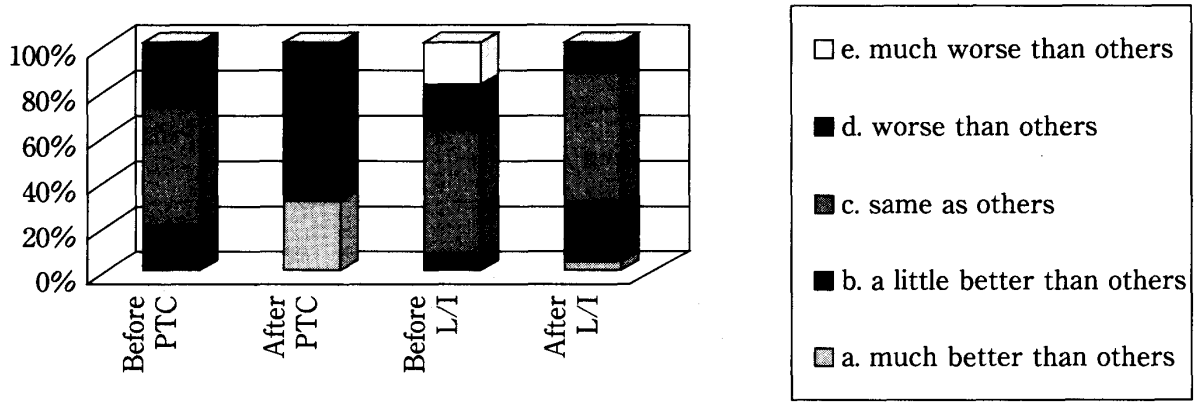
The results and analysis of test scores section must begin with a disclaimer. Due to the absence of baseline pre-test scores, the analysis of test scores is inconclusive. The results of test scores will be presented as a matter of interest only. The reader is advised that reliable correlations or cause and effect relationships between using the PTC and test scores cannot be made from this data. Also, in this section, three groups will be compared, the PTC IIa Group, the L/I IIa Group, and the small PTC 2X Group (a splinter group who studied two times per week outside of the IIa class).

As mentioned in the Methodology section, two types of examinations were administered during the last two weeks of the fall term. One was a speaking test where each student had a ten-minute period to converse with the PTC. The number of acceptable conversations that the PTC awarded the student was her score. The other test consisted of teacher-monitored conversations between two students with the teacher recording impressions. Students randomly selected the subject matter of six conversations (all taken from the *Let's Go To New York* software), then talked about the topics. They could

Table 4 : Machine Test and Teacher-monitored Conversation Test Results

Score	PTC Machine	L/I Machine	2X Week Machine	PTC Teacher	L/I Teacher	2X Week Teacher
Average	3.2	5	6.8	2.7	2.2	1.9
Median	2	4	7	2	2	2
Mode	1	1	9	1	2	1

Chart 2 : End of Term Self-assessment of Pronunciation Ability



comparison of pronunciation self-assessments.

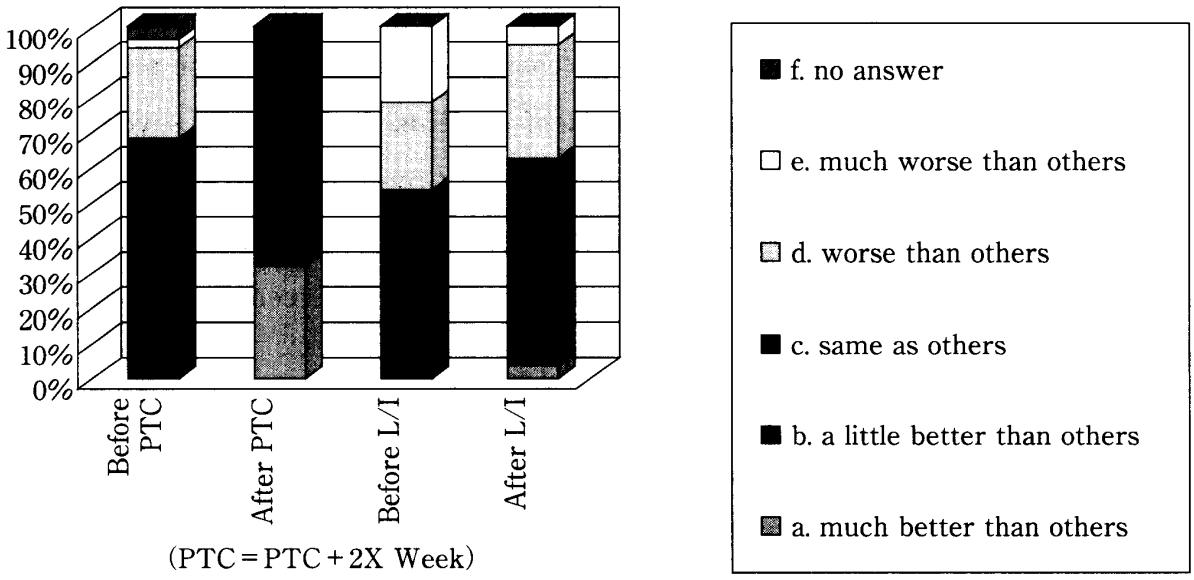
Prior to participating in the research project, when asked to assess whether a native speaker could understand their speaking, students in all groups predominantly responded *maybe* with a small percentage indicating *no*. At the conclusion of the research period, twenty percent of students in all groups responded that *yes* a native-speaker could understand their speaking.

In response to how students felt when speaking English to their native English-speaking teachers and their Japanese English teachers, before the research, the PTC Group showed significantly higher confidence than either the L/I Group or the 2X Week Group. This could be attributed to the fact that thirteen percent of the PTC Group had previously studied English at a language school and twenty-six percent had begun studying English in elementary school, therefore giving these individuals greater confidence in speaking English. However, after the research period, forty-two to forty-seven percent of students from both the PTC group and the L/I Group answered that they felt *more confident*.

Considering the questionnaire results for speaking and pronunciation from a statistical perspective, four two-sample t-tests assuming equal variance were conducted. The results of the PTC Group and L/I Group before study and after study were compared for both speaking and pronunciation ability self-assessment. At a confidence level of .05, a significant difference for speaking ability was found, but not for pronunciation. This indicates that the PTC has a positive impact on student's self-perception of their speaking

cent *better than other students*, thirty-four percent the *same as other students*, but still thirty-four percent *worse than other students* and one percent *much worse than other students*. Chart 1 illustrates these differences.

Chart 1 : End of Term Self-assessment of Speaking Ability



At the end of the term after using the PTC, thirty-three percent of the students considered that their pronunciation had improved to be *much better than other students*, and an additional fourteen percent felt that they improved to be *a little better than others*; however, an additional eleven percent joined the original twenty-two percent who thought their pronunciation was *worse than other students*. Since the PTC does not differentiate between students or show sympathy to students who are trying but do not achieve recognizable results, it seems plausible that some students might feel that they had regressed rather than progressed. This high result is a concern.

Showing the same trend as for the speaking results, in the L/I Group, only one percent felt that they were *much better than other students* while the section a *little better than other students* jumped to twenty-six percent. The *same as others* percentage remained the same at fifty-eight percent, but the *worse than others* category shrank by six percent to eleven percent and no students believed themselves to be *much worse than others* after practicing the L/I method with the teacher for the term. Chart 2 illustrates the

register for the elective communication class, that the class size was small and students felt more confident, or that these were the same students who in the PTC Group responded that they were a *little better at speaking than others*. Since responses were anonymous, it is not possible to track or explain the exact reason.

As for pronunciation, Table 2 contains data collected at the beginning of the term. No students thought their pronunciation was *better than other students* while a striking forty percent of the 2X Week Group members considered themselves to be a *little better than other students*. As anticipated, the bulk of students, forty to fifty-seven percent, rated themselves the *same as other students*. The L/I Group reported a markedly lower self-perception of their pronunciation abilities than the other two groups with one percent responding *a little better than other students* and seventeen percent *much worse than other students*.

Table 2 : Beginning of Term Self-assessment of Pronunciation Ability

Compared to other Wayo students, how is your pronunciation?	PTC	L/I	2X Week
a. much better than other students	0.0	0.0	0.0
b. a little better than other students	13.0	1.0	40.0
c. the same as other students	57.0	58.0	40.0
d. worse than other students	22.0	17.0	20.0
e. much worse than other students	0.5	17.0	0.0

At the end of the research period, all groups thought their speaking and pronunciation abilities had improved, with the PTC and 2X Week groups showing the greatest perceived improvement. For example, thirty-three percent of the groups who practiced with the PTC thought their speaking was *much better than other students*, thirty-three percent thought it was *better than other students*, thirty-three percent thought they were the *same*, and none of them thought they were *worse or much worse than other students*.

Conversely, the L/I Group's self-assessment of improvement did not reach such high levels. Only one percent responded *much better than other students*, twenty-nine per-

or worse than other students in the group, however, the results were not adjusted in anyway to allow for this possibility. Responses represent percentages of the whole group as recorded.

In response to questions about nationality, living experience in an English-speaking country, number of years and methods of studying English, answers were virtually identical. All students were Japanese and none had ever lived abroad. All had studied English for more than six years in junior and senior high school and thirty-seven to fifty-two percent had studied English at a *juku* prior to entering university. Thirty percent of the PTC Group and forty percent of the 2X per Week Group had also studied English as a language school compared to thirteen percent of the L/I Group. Also, twenty-six percent of the PTC Group had started studying English in elementary school compared with no students in the other groups.

For purposes of this article, only the most pertinent questions from this questionnaire were analyzed. Although questions pertaining to reading and listening were also asked, the results did not contribute useful data for the analysis undertaken here. Therefore, questions regarding speaking and pronunciation ability and attitudes about oral communication in English were used.

As we can see from the beginning of the term self-assessment responses contained in Table 1, most students believed that their English speaking ability was *the same or worse than* other students at Wayo with the exception of the 2X Week Group where forty percent responded *a little better than other students*. This higher response rating could be due to the fact that the students in this class liked or wanted to study English enough to

Table 1 : Beginning of Term Self-assessment of Speaking Ability

Compared to other Wayo students, how well can you speak English?	PTC Group	L/I Group	2X Week
a. much better than other students	0.0	0.0	0.0
b. a little better than other students	1.0	0.5	40.0
c. the same as other students	57.0	29.0	40.0
d. worse than other students	26.0	25.0	20.0
e. much worse than other students	2.0	21.0	0.0

viously said. A small optical window showing the attempted utterance appeared and the sections that were not comprehensible were highlighted. The student had three or four chances to repeat the sentence correctly. If at the end of that time the utterance still was not understandable, the computer proceeded to the next part of the conversation. When this occurred, the student lost the chance to obtain a heart for successful completion of the conversation.

The L/I Group students practiced the same *Let's Go To New York* conversations from prints. Students listened to the teacher read the conversation and repeated line by line. Next, students repeated the entire conversation while the teacher listened, suggesting pronunciation, rhythm, or intonation improvements, letting students correct their errors, and explaining the meaning of words or expressions. Finally, students practiced the conversations four times with a student partner. While the teacher walked among them, listening and pointing out areas for improvement.

The 2X Week Group consisted of seven students (who were also members of either the PTC Group or the L/I Group). These students used the PTC for thirty minutes in their elective Communicative English IIb class each week. The difference in the practice sessions was that, due to the small class size, each student could usually use the equipment individually for a full twenty to thirty minutes without sharing with other students.

At the end of the term, all students' speaking was tested in two ways to determine their levels of speaking ability. One test was administered using the PTC and *Let's Go To New York* software. The teacher administered the second conversation test using the same topics as the machine test, but students spoke with another student. (Note: One week before the machine test, students in the L/I Group practiced with the PTC to prevent skewed results due to student unfamiliarity with the equipment.)

Results and Analysis of Student Self-assessment of Language Abilities

At the beginning of the term all students answered a questionnaire that provided information about their educational background with respect to studying English as well as self-perceptions of their abilities compared to their classmates and levels of confidence communicating in English. The researcher acknowledges the possibility of a cultural tendency for students to not want to stand out from the main group, i.e., either better than

English class by using the PTC than by practicing the teacher-directed listen and imitate (L/I) method. Additionally, a splinter group used the same equipment and software during one additional class each week. Comparisons with this group were also made. Test results and student perceptions of improvement were collected at the end of the term.

Methodology

Research began in October 1999 at Wayo Women's University. Participants in the study were fifty-five sophomore International Social Studies students, all of whom were placed into the required Communicative English IIa course in numerical order according to their student numbers, and a group of seven of the same students who had enrolled in an additional elective class, Communicative English IIb. In the Communicative English IIa classes, students were virtually randomly selected since no placement or ability ranking was applied in the selection. In the two Communicative English IIa classes, students were divided into two groups subsequently referred to as the PTC Group (twenty-six students) and L/I Group (twenty-nine students). Students in both groups answered a questionnaire at the beginning of the term and at the end of the term. The purpose of these questionnaires was to determine each student's self-perception of her own speaking ability and attitude towards using English to communicate. At the end of the term, all students were given both a machine conversation test and a teacher-monitored test.

Each week for thirty minutes students in the PTC Group practiced spoken conversations using the PTC ASR device and *Let's Go To New York* software. The class was divided into groups of four or five and each group was given a PTC to use. Each student also received a copy of the conversations contained on the PTC software and a tally sheet to record successfully completed conversations. While one student in the group used the PTC, others were supposed to read and practice the conversations individually. When using the PTC, students received a "heart" for each successfully completed conversation. They marked these hearts on the tally sheet. To successfully complete a conversation, the student conducted an oral conversation with the tiny computer. The computer asked questions or made comments to which the student responded. If the computer understood the student's response, the conversation continued. If it did not understand the utterance, it said, "Excuse me ? " and waited for the student to repeat or correct what she had pre-

Automatic Speech Recognition and Language Learning: Research Results

Rebecca Tanaka

Introduction

Instructional technology (IT) is one of the most talked about topics within many academic circles lately, particularly in the area of language learning. Proponents of IT tout the benefits of electronic equipment and language–learning software. For example, IT offers a more student–focused form of instruction, it allows students to work at their own pace, it replicates genuine communication, and *help options*, such as dictionaries, thesauruses, direct text translations, pronunciation guides, etc., are instantly available. New applications, hardware and software enter the market frequently.

Amidst all these attractive new developments, however, lies the question of whether or not IT can measurably improve student learning. This paper reviews one research attempt to determine the impact of one type of IT, an automatic speech recognition (ASR) device, on student learning. In–class research was conducted during the fall 1999 term and results of that research are presented in this article.

For background information on ASR technology, how it works, and a review of the various forms of ASR, refer to Tanaka, 1999.

Objectives of the Study

The primary objective of this study was to determine the effect of a particular type of ASR technology device, specifically the Pocket Talking Coach (PTC) manufactured by Fuji Xerox, on the speaking ability of a defined group of university students in Japan. The following hypothesis was tested: the English speaking ability of Japanese university students improves more in one semester in a typical 90–minute once–per–week university